Science of the Total Environment 825 (2022) 153982

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Prediction of enteric methane production and yield in dairy cattle using a Latin America and Caribbean database

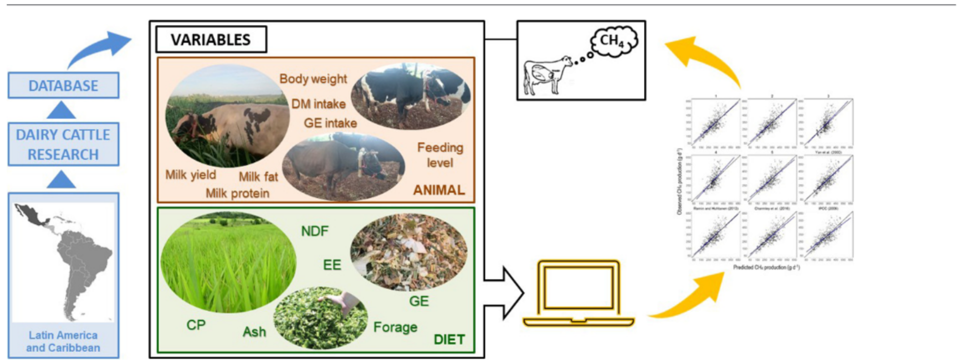
Guilhermo F.S. Congio [a,*], André Bannink [b], Olga L. Mayorga [c], João P.P. Rodrigues [d], Adeline Bougouin [e], Ermias Kebreab [e], Ricardo R. Silva [f], Rogério M. Maurício [g], Sila C. da Silva [a], Patrícia P.A. Oliveira [h], Camila Muñoz [i], Luiz G.R. Pereira [j], Carlos Gómez [k], Claudia Ariza-Nieto [c], Henrique M.N. Ribeiro-Filho [l], Octavio A. Castelán-Ortega [m], Jaime R. Rosero-Noguera [n], Maria P. Tieri [o,p], Paulo H.M. Rodrigues [q], Marcos I. Marcondes [r], Laura Astigarraga [s], Sergio Abarca [t], Alexander N. Hristov [u]

[a] Department of Animal Science, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, SP 13418-900, Brazil
[b] Wageningen Livestock Research, Wageningen University & Research, Wageningen, AH 6700, the Netherlands
[c] Colombian Corporation for Agricultural Research, Tibaitatá, Bogotá D.C. 250047, Colombia
[d] Faculty of Animal Science, Federal University of Southern and Southeastern Pará, Xinguara, PA 68555-110, Brazil
[e] Department of Animal Science, University of California, Davis, CA 95618, USA
[f] Department of Animal Science, Federal University of Minas Gerais, Belo Horizonte, MG 31270-901, Brazil
[g] Department of Bioengineering, Federal University of São João del-Rei, São João del-Rei, MG 36307-352, Brazil
[h] Brazilian Agricultural Research Corporation, Embrapa Southeast Livestock, São Carlos, SP 13560-970, Brazil
[i] Instituto de Investigaciones Agropecuarias, INIA Remehue, Osorno 5290000, Chile
[j] Brazilian Agricultural Research Corporation, Embrapa Dairy Cattle, Juiz de Fora, MG 36038-330, Brazil
[k] Department of Animal Husbandry, Faculty of Animal Science, National Agrarian University La Molina, Lima 15024, Peru
[l] Department of Animal and Food Science, Santa Catarina State University, Lages, SC 88520-000, Brazil
[m] Faculty of Veterinary Medicine and Animal Science, Autonomous University of the State of Mexico, Toluca, Estado de México 5000, Mexico
[n] Faculty of Agricultural Sciences, University of Antioquia, Medellín, Antioquia 050034, Colombia
[o] National Institute of Agricultural Technology, Rafaela, Santa Fé S2300, Argentina
[p] Regional Faculty of Rafaela, National Technological University, Rafaela, Santa Fé S2300, Argentina
[q] Department of Animal Nutrition and Production, Faculty of Veterinary Medicine and Animal Science, University of São Paulo, Pirassununga, SP 13635-900, Brazil
[r] Department of Animal Sciences, Washington State University, Pullman, WA 99163, USA
[s] Department of Animal Science and Pastures, Faculty of Agronomy, University of the Republic of Uruguay, Montevideo 12900, Uruguay
[t] National Institute of Innovation and Agricultural Technology Transfer, Turrialba, Cartago 30508, Costa Rica
[u] Department of Animal Science, The Pennsylvania State University, University Park, PA 16802, USA

HIGHLIGHTS

- Dry matter intake (DMI) was the most important predictor of dairy CH_4 production.
- Simple regression models including DMI were accurate for predicting CH_4 production.
- CH_4 production can also be predicted using milk yield when DMI is missing.
- Developed models outperformed IPCC Tier 2 equations.
- These newly-developed models can improve the accuracy GHG inventories from LAC countries.

GRAPHICAL ABSTRACT



Abbreviations: BIC, Bayesian information criterion; BW, body weight; CCC, concordance correlation coefficient; CP, crude protein; DM, dry matter; DMI, DM intake; EE, ether extract; EPCM, energy and protein-corrected milk; FL, feeding level; For, forage; GE, gross energy; GEI, GE intake; GHG, greenhouse gas; IPCC, Intergovernmental Panel on Climate Change; LAC, Latin America and Caribbean; MB, mean bias; MF, milk fat; MP, milk protein; MSPE, mean square prediction error; MY, milk yield; NDF, neutral-detergent fiber; RMSPE, root MSPE; RSR, RMSPE to standard deviation of observed values ratio; SD, standard deviation; SB, slope bias; VIF, variance inflation factor.

* Corresponding author at: Department of Animal Science, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba 13418-900, SP, Brazil.
E-mail address: gcongio@gmail.com (G.F.S. Congio).

http://dx.doi.org/10.1016/j.scitotenv.2022.153982

A R T I C L E I N F O

Article history:
Received 13 January 2022
Received in revised form 8 February 2022
Accepted 14 February 2022
Available online 22 February 2022

Editor: Ouyang Wei

Keywords:
Diet
Empirical modeling
Enteric methane
GHG inventory
Linear models
Prediction equations

A B S T R A C T

Successful mitigation efforts entail accurate estimation of on-farm emission and prediction models can be an alternative to current laborious and costly *in vivo* CH_4 measurement techniques. This study aimed to: (1) collate a database of individual dairy cattle CH_4 emission data from studies conducted in the Latin America and Caribbean (LAC) region; (2) identify key variables for predicting CH_4 production (g d^{-1}) and yield [g kg^{-1} of dry matter intake (DMI)]; (3) develop and cross-validate these newly-developed models; and (4) compare models' predictive ability with equations currently used to support national greenhouse gas (GHG) inventories. A total of 42 studies including 1327 individual dairy cattle records were collated. After removing outliers, the final database retained 34 studies and 610 animal records. Production and yield of CH_4 were predicted by fitting mixed-effects models with a random effect of study. Evaluation of developed models and fourteen extant equations was assessed on all-data, confined, and grazing cows subsets. Feed intake was the most important predictor of CH_4 production. Our best-developed CH_4 production models outperformed Tier 2 equations from the Intergovernmental Panel on Climate Change (IPCC) in the all-data and grazing subsets, whereas they had similar performance for confined animals. Developed CH_4 production models that include milk yield can be accurate and useful when feed intake is missing. Some extant equations had similar predictive performance to our best-developed models and can be an option for predicting CH_4 production from LAC dairy cows. Extant equations were not accurate in predicting CH_4 yield. The use of the newly-developed models rather than extant equations based on energy conversion factors, as applied by the IPCC, can substantially improve the accuracy of GHG inventories in LAC countries.

1. Introduction

Methane (CH_4) is a powerful short-lived climate forcer (IPCC, 2018) and decreasing its emission is crucially important for limiting global warming to 2.0 °C above pre-industrial levels as defined in the Paris Agreement (UN General Assembly, 2015; Arndt et al., 2021). Successful mitigation efforts entail accurate estimation of on-farm emission (Niu et al., 2018). Accurate estimation of CH_4 emissions is also necessary to enable governments to improve their greenhouse gas (GHG) inventories, which is the foundation for policy makers to develop mitigation plans (Moraes et al., 2014).

Several empirical prediction models have been developed (Niu et al., 2018; van Lingen et al., 2019) as an alternative to current *in vivo* CH_4 measurement techniques (Hammond et al., 2016). These models can be useful to estimate enteric CH_4 emissions without undertaking extensive and costly experiments (Patra and Lalhriatpuii, 2016; Hristov et al., 2018). Recent meta-analyses, however, were based primarily on data from the U.S. and the E.U. with no or minimal data from the Latin America and Caribbean (LAC) region (Moraes et al., 2014; Niu et al., 2018; van Lingen et al., 2019). These analyses concluded that region-specific models are more accurate in predicting enteric CH_4 production than global models, mainly due to specifics regarding animal diets and feed management systems (Niu et al., 2018; van Lingen et al., 2019).

Dairy cattle in the LAC region emitted 54 MT of CO_2 equivalents from enteric CH_4 fermentation, comprising 14.3% of global dairy cattle emissions in 2018 (FAOSTAT, 2020). Two CH_4 modeling studies were recently conducted using databases from LAC countries (Benaouda et al., 2020; Ribeiro et al., 2020). These studies, however, used limited databases which resulted in models with moderate accuracy and restricted use given the wide diversity of dairy production systems found in the region. In this context, the objectives of the current study were to: (1) collate a database of individual dairy cattle enteric CH_4 emission data from studies conducted in the LAC region; (2) identify key dietary and animal variables for predicting enteric CH_4 production (g d^{-1}) and yield [g kg^{-1} of feed dry matter intake (DMI)]; (3) develop and cross-validate these newly-developed models; and (4) compare their predictive ability with extant equations, including both from IPCC (1997, 2006), which are currently used to support national GHG inventories in the LAC region.

2. Material and methods

2.1. Database

The LAC methane project is an international collaborative initiative specifically designed to involve animal scientists from the LAC region (Congio et al., 2021). The resultant dairy cattle CH_4 database collated in the frame of the LAC methane project included 1327 individual dairy cattle records from 42 published (n = 15) and unpublished (n = 27) studies conducted from 2012 to 2021 by researchers from eight countries in the LAC region (Brazil, n = 788 records from 20 studies; Costa Rica, n = 182 from 2 studies; Colombia, n = 135 from 9 studies; Chile, n = 81 from 2 studies; Peru, n = 57 from 3 studies; Argentina, n = 36 from 1 study; Mexico, n = 32 from 4 studies; and Uruguay, n = 16 from 1 study). The database comprised records of enteric CH_4 production along with corresponding DMI, body weight (BW), dietary contents of neutral-detergent fiber (NDF), ether extract (EE), crude protein (CP), ash, gross energy (GE) and forage. It also included milk yield (MY), and contents of milk fat (MF) and milk (crude) protein (MP). Studies containing missing dietary parameters were completed according to the literature as follows. Dietary GE was calculated (n = 327) based on an equation derived from Weiss and Tebbe (2019): GE (MJ kg^{-1} DM) = {[(CP (%) × 0.056) + (EE (%) × 0.094)] + [(100 – CP – EE – ash (%)) × 0.042] × 4.184}. For estimating dietary EE (n = 121), local literature was used according to each study. When MY and its composition were known, the energy and protein-corrected milk (EPCM) was calculated according to NRC (2001): EPCM (kg d^{-1}) = [(0.327 × kg of milk yield) + (12.95 × kg of fat yield) + (7.20 × kg of protein yield)]. Feeding level [FL = DMI (kg) ÷ BW (kg) × 100], CH_4 yield [CH_4 yield (g kg^{-1}) = CH_4 production (g d^{-1}) ÷ DMI (kg d^{-1})], GE intake [GEI = DMI (kg) × dietary GE (MJ kg^{-1})] and CH_4 conversion factor [Y_m; CH_4 production (g d^{-1}) × 0.05565 ÷ GEI (MJ d^{-1}) × 100)] were calculated for all records.

Records with missing CH_4 yield were removed from the database (n = 135). Data from each study were graphically evaluated and eight studies (n = 292) were removed due to negative relationships between CH_4 production and DMI (Hristov et al., 2018). In addition, treatments containing inclusion levels of feed supplements and additives with known anti-methanogenic effects (*e.g.*, lipid supplementation, monensin, tannin-extracts) were also removed (n = 152). Other outliers were screened using the interquartile range method (Zwillinger and Kokoska, 2000) based on CH_4 production and yield, Y_m, DMI, GEI, BW, MY, EPCM, MF, and MP. A factor of 1.5 for extremes was used in constructing markers to identify outliers (Niu et al., 2018). After removing all outliers, the final database retained 34 studies and 610 individual animal records, 46% of the original database. Niu et al. (2018) and van Lingen et al. (2019), using a similar approach for dairy and beef cattle, retained 49 and 50%, respectively, of the observations in their original databases. The complete data bibliography of the final database is provided in Supplementary Material.

2.2. Model development

Model development was performed using a sequential approach by incrementally adding different levels of variables to develop models with increasing complexity (Niu et al., 2018). For CH_4 production, the first step included simple regression models based on DMI, GEI, MY, or EPCM. In the second step, multiple regression models tested combinations of DMI or GEI with FL, CP, EE, or NDF separately. Both MF and MP were tested with MY or EPCM. Lastly, individual models were tested by combining: a selection of dietary parameters, a selection of DMI or GEI with all dietary parameters, a selection of all available variables, and a selection of all available variables except DMI or GEI. Methane yield regression models were developed without DMI as predictor because this variable already has been used for the calculation of reported CH_4 yield (Niu et al., 2018). Simple CH_4 yield regression models were based on BW, FL, dietary contents of NDF, EE, and CP, MY, and EPCM. In the second step, multiple regression models tested combinations of MY or EPCM with milk composition parameters. Then, additional multiple CH_4 yield regression models tested a selection of dietary parameters and a selection of all the available variables except DMI or GEI.

Production and yield of CH_4 were predicted by fitting mixed-effects models using lme4 (Bates et al., 2015) package of R statistical software (R Core Team, 2020; version 4.0.2) according to the following equation:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \ldots + \beta_k X_{ijk} + S_i + \varepsilon_{ij}$$

where Y_{ij} is the response variable of CH_4 production (g d^{-1}) or CH_4 yield (g kg^{-1} DMI); β_0 is the fixed effect of intercept; X_{ij1} to X_{ijk} are the fixed effects of predictor variables and β_1 to β_k are the corresponding slopes; S_i and ε_{ij} are the random effect of study and residual error, respectively. Covariates that play a key role in predicting CH_4 variables were selected for multiple regression models using the backward multistep selection approach (van Lingen et al., 2019). The Bayesian information criterion (BIC) was computed, and models with the smallest BIC were selected because smaller BIC indicates a better tradeoff between the goodness of fit and model complexity. Additionally, the presence of multicollinearity of fitted models was checked using the variance inflation factor (VIF), and models were selected only if all predictor variables had a VIF lower than 3 (Zuur et al., 2010).

2.3. Cross-validation and model evaluation

The predictive accuracy of fitted CH_4 prediction models was evaluated using a leave-one-out cross-validation (James et al., 2014). Studies were considered as folds and, in each simulation, one study was removed as a testing set and the remaining were used as a training set (Ribeiro et al., 2020). The predictions of all folds were used to conduct the model evaluation. Equations based on energy conversion factors from IPCC (1997, 2006), currently used to support national GHG inventories in the LAC region, were evaluated. Furthermore, extant equations from Yan et al. (2000), Ellis et al. (2007), Hristov et al. (2013), Nielsen et al. (2013), Ramin and Huhtanen (2013), Moraes et al. (2014), Storlein et al. (2014), Charmley et al. (2016), Patra (2017), Niu et al. (2018), Benaouda et al. (2020), and Ribeiro et al. (2020) were also evaluated. The best-performing equation from each study was selected based on the availability of predictors in the current database. Data from studies in the current database used to develop the above extant equations were excluded from evaluations of those extant equations to ensure independent evaluation (van Lingen et al., 2019). Evaluation of developed models and extant equations was assessed on complete (all-data), confined, and grazing cows subsets.

A combination of metrics was used to assess model performance. The mean square prediction error (MSPE) was calculated according to Bibby and Toutenburg (1977) as:

$$MSPE = \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{n}$$

where O_i is the observed value of the response variable for the i^{th} observation, P_i is the predicted value of the response variable for the i^{th} observation, and n is the number of observations. The root MSPE (RMSPE) was calculated and used to assess overall model prediction accuracy. It was expressed as a proportion of observed CH_4 production or yield means, and smaller RMSPE indicates better model performance. The RMSPE to standard deviation of observed values ratio (RSR), used to assess the specific variability of the data used for evaluation (Moriasi et al., 2007), was calculated as:

$$RSR = \frac{RMSPE}{S_o}$$

where S_o is the standard deviation (SD) of observed values. Smaller RSR indicates less variation in the prediction error relative to the standard deviation of the observed values. In the current analysis, we considered unsuitable models that presented RMSPE greater than the SD of observed values (RSR $\geq$ 1.00) (van Lingen et al., 2019). Additionally, the MSPE was decomposed into sources of errors including mean bias (MB) and slope bias (SB), measures of precision and accuracy, respectively (Bibby and Toutenburg, 1977), of which were calculated as:

$$MB = \left(\overline{P} - \overline{O}\right)^2$$

$$SB = \left(S_p - r \times S_o\right)^2$$

where $\overline{P}$ and $\overline{O}$ are the predicted and observed CH_4 parameter means, S_p is the SD of the predicted values, and r is the Pearson correlation coefficient. Finally, the concordance correlation coefficient (CCC; Lin, 1989) was calculated as follows:

$$CCC = r \times C_b$$

where C_b is the bias correction factor. It is a metric that accounts for precision and accuracy, and values closer to 1 indicate better model performance.

3. Results

3.1. Database

Summary statistics for all-data, confined, and grazing subsets that included DMI, BW, FL, dietary nutrient composition, milk parameters, and CH_4 emission variables are shown in Table 1. Overall, the all-data subset was mostly comprised of confined rather than grazing animals (330 *vs.* 280 records). The confined subset was composed of 80% lactating and 20% dry cows, whereas the grazing subset included 99% of lactating dairy cows. Confined animals were heavier and comprised mostly Holstein × Gyr (49%), Holstein (27%), and Gyr (16%). Grazing dairy cows were predominantly Holstein × Jersey (40%), Holstein (33%), and Brown Swiss (8%). The main forage types for confinement systems were corn silage (58%), corn silage plus tropical hays (16%), fresh-cut forage (*Pennisetum clandestinum* or *Saccharum officinarum*, 10%), and corn silage plus temperate hays (8%). Under grazing, typical pastures were composed by *Pennisetum spp.* (31%), *Lolium spp.* (21%), *Megathyrsus maximus* (17%), *Urochloa spp.* (16%), *Medicago sativa* (8%), or *Festuca spp.* (6%). Forage and NDF contents were markedly higher for grazing than confined dairy cows. Additional particularities were observed between subsets. Methane emissions were estimated primarily using respiration chambers (60%) and secondly through sulfur hexafluoride (SF_6; 39%) under confinement, whereas SF_6 (86%) was the most used technique under grazing. Still, DMI was estimated gravimetrically (100%) and using markers (89%) in confinement and grazing systems, respectively.

Table 1
Summary statistics for variables used in the analysis of all-data, confined, and grazing subsets of the Latin America and Caribbean dairy cattle database.

Item[a]	All-data					Confined					Grazing				
	n[b]	Mean	Min[b]	Max[b]	SD[b]	n	Mean	Min	Max	SD	n	Mean	Min	Max	SD
DMI (kg d^{-1})	610	14.7	4.50	25.2	4.52	330	14.8	4.50	25.2	4.78	280	14.5	5.64	24.1	4.21
GEI (MJ d^{-1})	610	262	85.0	445	82.5	330	268	85.0	445	87.3	280	255	95.0	427	75.9
BW (kg)	610	520	291	1021	91.4	330	542	352	1021	91.2	280	494	291	694	84.7
FL (DMI as % BW)	610	2.82	0.97	5.19	0.770	330	2.73	0.97	4.59	0.765	280	2.93	1.07	5.19	0.762
Diet composition (% DM)															
NDF	610	41.4	16.1	67.7	9.75	330	37.3	22.6	60.0	8.01	280	46.3	16.1	67.7	9.40
EE	610	3.00	1.40	6.69	0.791	330	3.09	1.40	6.69	0.937	280	2.89	1.61	4.25	0.556
CP	610	16.0	7.20	24.9	2.76	330	15.8	10.5	20.2	2.40	280	16.3	7.20	24.9	3.10
Ash	610	8.07	3.90	16.6	2.01	330	7.43	4.50	12.7	1.82	280	8.83	3.90	16.6	1.96
GE (MJ kg^{-1} DM)	610	17.7	15.2	19.3	0.596	330	17.9	15.2	18.9	0.670	280	17.5	16.6	19.3	0.423
Forage	610	67.9	8.1	100	16.7	330	63.0	43.8	94.0	11.0	280	73.6	8.13	100	20.1
Yield (kg d^{-1})															
MY	539	18.3	1.50	40.1	7.55	263	18.6	4.51	37.8	6.97	276	18.0	1.50	40.1	8.06
EPCM	487	19.8	2.36	41.1	6.75	247	20.0	5.15	33.8	6.49	240	19.6	2.36	41.1	7.01
Milk composition (%)															
MF	487	3.88	1.60	7.21	0.817	247	4.18	1.83	6.57	0.728	240	3.58	1.60	7.21	0.793
MP	487	3.27	2.30	4.96	0.373	247	3.33	2.47	4.44	0.364	240	3.22	2.30	4.96	0.375
Methane emissions															
CH_4 (g d^{-1})	610	309	86.9	612	98.8	330	308	86.9	594	108	280	309	152	612	86.9
CH_4 per DMI (g kg^{-1})	610	21.5	11.9	41.5	4.45	330	20.8	13.1	30.2	2.99	280	22.4	11.9	41.5	5.60
Y_m (% GEI)	610	6.71	3.71	12.9	1.43	330	6.40	4.11	9.25	0.902	280	7.08	3.71	12.9	1.81

[a] DMI = dry matter intake; GEI = gross energy intake; BW = body weight; FL = feeding level; NDF = dietary neutral-detergent fiber; EE = dietary ether extract; CP = dietary crude protein; GE = dietary gross energy; MY = milk yield; EPCM = energy and protein-corrected milk; MF = milk fat; MP = milk protein; Y_m = methane conversion factor.
[b] n = number of observations; Min = minimum; Max = maximum; SD = standard deviation.

3.2. Methane production models

Methane production prediction equations and model performance indicators are presented in Table 2. Dry matter intake (Eqs. 1, 6, and 11), GEI (Eqs. 2, 7, and 12), MY (Eqs. 3, 8, and 13), and EPCM (Eqs. 4, 9, and 14) indicated a positive relationship with CH_4 production. Overall, multiple regression models that included dietary parameters and DMI or GEI (equations not shown) did not increase the predictive ability compared with DMI and GEI simple regressions models in all subsets. Still, the inclusion of MF and MP did not improve MY or EPCM simple regression models (equations not shown). Multiple regression models including only dietary parameters had the worst predictive performance among developed equations in all subsets (equations not shown).

Simple regression models developed on all-data subset including either DMI (Eq. 1) or GEI (Eq. 2) were of comparable accuracy with negligible systematic biases (Fig. 1). Still, one-variable models including either MY

Table 2
Enteric CH_4 production (g d^{-1}) prediction equations and model performance for the all-data, confined, and grazing subsets of the Latin America and Caribbean dairy cattle database.

Subset Equation	Prediction equation[a]	Model performance[b] n[b]	RSR	RMSPE, %	MB, %	SB, %	CCC
All-data							
(1)	40.7 (11.8) + 18.0 (0.592) × DMI	610	0.64	20.5	0.56	0.64	0.76
(2)	42.1 (11.7) + 1.00 (0.033) × GEI	610	0.63	20.2	0.29	0.86	0.77
(3)	178 (136) + 7.21 (0.547) × MY	539	0.79	23.1	0.83	0.49	0.53
(4)	153 (14.5) + 8.26 (0.547) × EPCM	487	0.76	22.2	0.49	1.56	0.57
(5)	30.6 (14.1) + 16.3 (0.838) × DMI + 2.04 (0.522) × EPCM	487	0.67	19.7	0.17	1.98	0.74
Confined							
(6)	4.28 (12.8) + 19.8 (0.686) × DMI	330	0.45	15.9	4.45	0.38	0.88
(7)	7.91 (12.1) + 1.09 (0.037) × GEI	330	0.43	15.1	2.01	0.41	0.90
(8)	157 (20.3) + 8.49 (0.745) × MY	263	0.79	23.6	1.31	0.60	0.53
(9)	130 (19.5) + 9.67 (0.682) × EPCM	247	0.71	21.6	0.46	2.75	0.62
(10)	− 642 (235) + 20.4 (0.741) × DMI + 34.9 (13.4) × GE + 5.92 (4.64) × EE − 4.06 (4.18) × MF + 3.53 (8.90) × MP	247	0.45	13.7	1.83	0.30	0.89
Grazing							
(11)	89.3 (20.4) + 15.7 (1.00) × DMI	280	0.91	25.6	0.83	9.16	0.48
(12)	87.9 (20.9) + 0.892 (0.058) × GEI	280	0.91	25.4	0.55	9.38	0.49
(13)	203 (17.4) + 5.70 (0.789) × MY	276	0.87	24.7	1.00	0.26	0.37
(14)	185 (20.4) + 6.39 (0.855) × EPCM	240	0.89	25.3	1.15	0.17	0.32
(15)	- 66.7 (446) + 14.2 (1.18) × DMI + 1.65 (25.5) × GE - 1.59 (17.6) × EE + 4.18 (4.26) × ash + 3.06 (0.808) × EPCM +16.0 (10.6) × MP	240	0.95	26.8	0.10	11.31	0.44

[a] DMI = dry matter intake (kg d^{-1}); GEI = gross energy intake (MJ d^{-1}); MY = milk yield (kg d^{-1}); EPCM = energy and protein-corrected milk (kg d^{-1}); GE = dietary gross energy (MJ kg^{-1} DM); EE = dietary ether extract (% DM); MF = milk fat (%); MP = milk protein (%); ash = dietary ash (% DM).
[b] n = number of observations used to fit equations and for model evaluation; RSR = RMSPE-observations standard deviation ratio; RMSPE = root mean square prediction error (% observed CH_4 production means); MB = mean bias (% MSPE); SB = slope bias (% MSPE); CCC = concordance correlation coefficient.

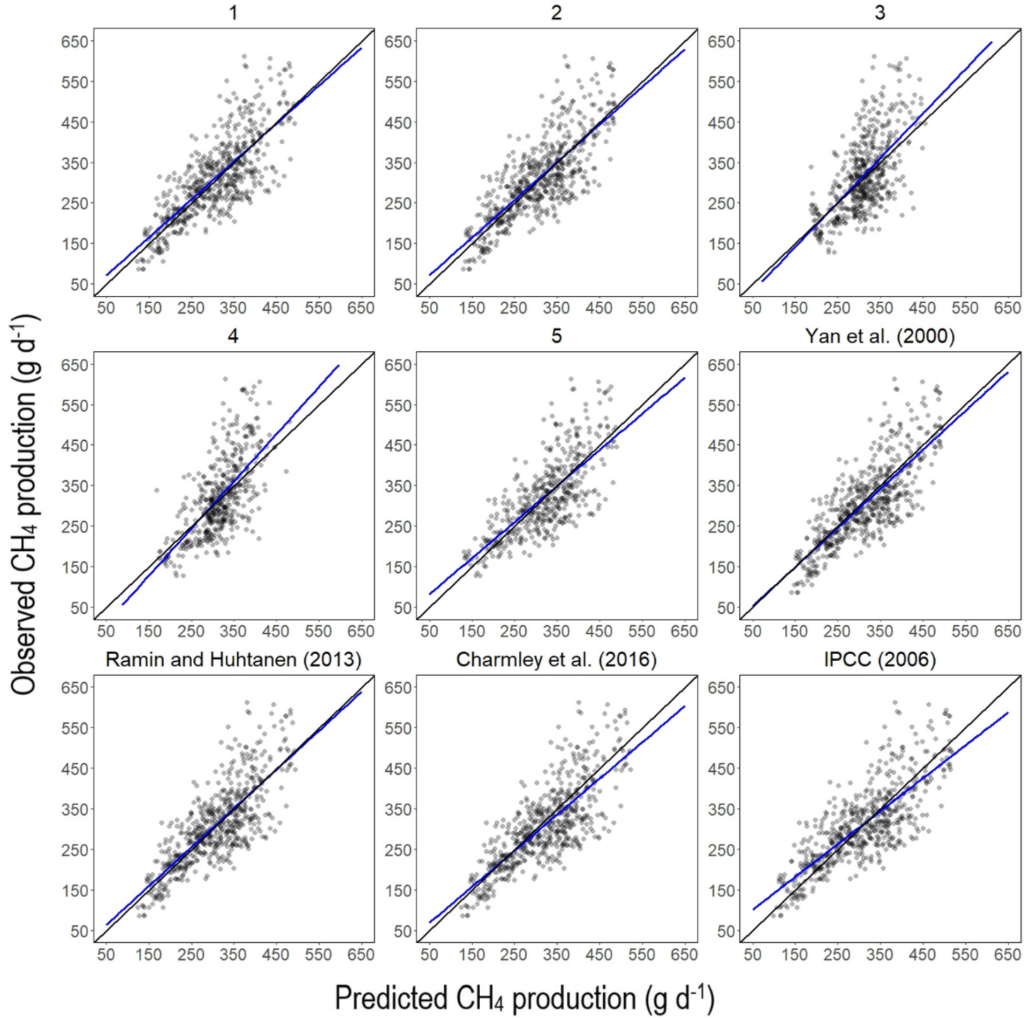


Fig. 1. Observed *vs.* predicted plots for all-data methane production (g d^{-1}) prediction equations. Developed models and extant equations are in accordance with Tables 2 and 3, respectively. The blue and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line (y = x), respectively.

(Eq. 3) or EPCM (Eq. 4) also had negligible systematic biases but were associated with larger RMSPE and RSR as well as smaller CCC compared to models 1 and 2. Model 5, which included DMI and EPCM, had the smallest RMSPE among all-data models.

For the confined subset, the simple regression model containing GEI (Eq. 7) performed slightly better than the DMI model (Eq. 6), with smaller RMSPE and RSR and larger CCC. Model 9, based on EPCM, had better performance with smaller RMSPE and RSR and larger CCC than the MY simple model (Eq. 8). The multiple regression model 10 presented the smallest RMSPE with negligible systematic biases (Fig. 2).

In accordance with the all-data subset, simple regression models including either DMI (Eq. 11) or GEI (Eq. 12) had similar overall performance in the grazing subset. Both models, however, had large SB (Fig. 3). The MY simple regression (Eq. 13) had the smallest RMSPE and RSR and negligible systematic biases, followed by model 14, which included EPCM. The multiple regression model 15 performed slightly worse than those previous simple models from the grazing subset.

Performance of extant equations for predicting enteric CH_4 production using all-data, confined, and grazing subsets are respectively shown in Tables 3, 4, and 5. Equations were ranked by RSR, which is the most appropriate statistic for evaluating equations based on different numbers of observations. Overall, simple equations based on feed intake (*i.e.*, DMI or GEI) had smaller RSR for predicting enteric CH_4 production. Equations from both Ramin and Huhtanen (2013) and Yan et al. (2000) had the smallest RSR and RMSPE with negligible systematic biases among all-data equations (Table 3). The equations by Charmley et al. (2016) and IPCC (2006) were the third and fourth-ranked RSR with low RMSPE. The IPCC (2006) equation over-predicted CH_4 at the high end and under-predicted it at the low end of production (Fig. 1). From the fifth to the fourteenth ranked RSR, equations were associated with a large and increasing MB (Fig. S1). Multiple regression equations from Ribeiro et al. (2020), Benaouda et al. (2020), and Niu et al. (2018), which included dietary parameters, had a low predictive ability and were outperformed by simpler equations in the all-data subset (Table 3).

For the confined subset, the IPCC (2006) Tier 2 equation outperformed all equations presenting the smallest RMSPE and RSR and the largest CCC, with negligible systematic biases (Table 4). Equations from Ramin and Huhtanen (2013), Yan et al. (2000), IPCC (1997), and Charmley et al. (2016) had similar predictive performance but were associated with either large MB or SB (Fig. 2). As observed in the all-data subset, multiple regression equations also had worse overall performance than simpler models in the confined subset (Fig. S2).

Both equations from Yan et al. (2000) and Ramin and Huhtanen (2013) had the best predictive performance among all extant equations in the grazing subset (Table 5). The multiple regression equation from Ribeiro et al. (2020), including GEI, BW, and dietary EE, was the third-ranked equation

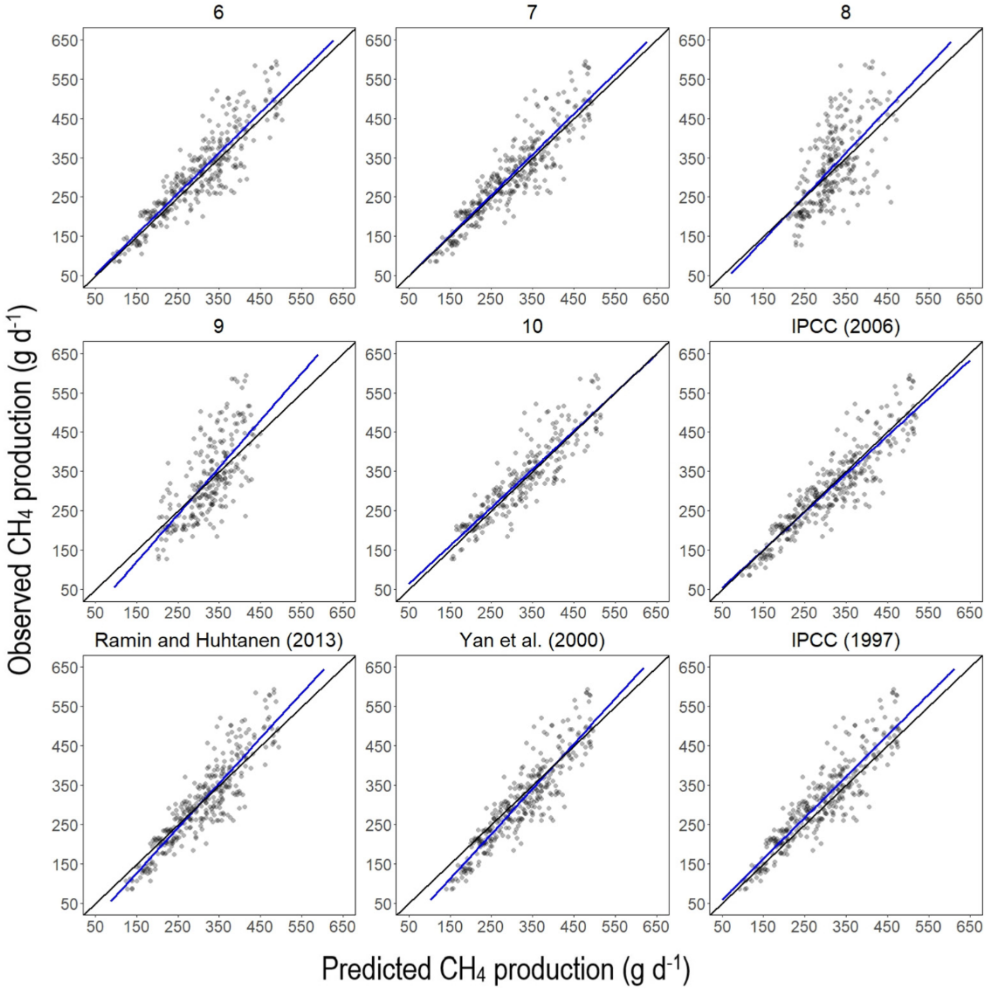


Fig. 2. Observed *vs.* predicted plots for confined cows methane production (g d^{-1}) prediction equations. Developed models and extant equations are in accordance with Tables 2 and 4, respectively. The blue and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line (y = x), respectively.

but had a large MB. The equation from Charmley et al. (2016) was the fourth-ranked RSR but was associated with large SB (Fig. 3). Lastly, equations from Moraes et al. (2014), Benaouda et al. (2020), Patra (2017), and Niu et al. (2018) had RSR > 1.00 (Fig. S3).

3.3. Methane yield models

Methane yield prediction equations and model performance indicators are shown in Table 6. Negative slope regression coefficients were obtained for BW (Eqs. 16, 17, and 19), FL (Eqs. 16, 18, and 19), whereas positive slope was observed for EPCM (Eq. 16) and dietary GE (Eq. 17). Considering all subsets, only the FL simple regression model (Eq. 18) developed on the grazing subset had RSR < 1.00. Multiple regression models including only dietary parameters as well as combining milk parameters had RSR > 1.00 in all subsets (equations not shown). Model 19, which included both BW and FL, had the smallest RSR and the largest CCC associated with negligible systematic biases. The CH_4 yield extant equation from Niu et al. (2018) presented RSR > 1.00 in all subsets (Table S1 and Fig. S4).

4. Discussion

Models predicting enteric CH_4 production in dairy cattle have been previously published (Nielsen et al., 2013; Moraes et al., 2014; Niu et al., 2018). These models, however, were mostly developed based on relatively small databases and/or focused on specific geographic regions that did not include LAC (*e.g.*, Ellis et al., 2007 comprised only studies from North America; Nielsen et al., 2013 included only studies from Nordic countries; Moraes et al., 2014 used only studies from one research station in the United States). Other studies have focused on tropical regions, but they were based on relatively small datasets (Benaouda et al., 2020; Ribeiro et al., 2020), and others using an intercontinental database that included minimal data from the LAC region (Niu et al., 2018). Previous research published by the 'Global Network' team reported that enteric CH_4 production is more accurately predicted by region- (Niu et al., 2018; van Lingen et al., 2019) and diet-specific (Benaouda et al., 2019) models; these authors indicated that additional efforts for important animal-producing regions are required. Our database includes the most available *in vivo* dairy cattle data regarding enteric CH_4 emission generated by researchers in the LAC region. Thus, this analysis is the most comprehensive effort to date to develop enteric CH_4 prediction models for dairy cattle managed under LAC conditions. An additional strength of the present study is the development of CH_4 yield models, whereas previous research focused primarily on CH_4 production.

Overall, CH_4 yield and Y_m averaged 21.5 g kg^{-1} DMI and 6.7% in the current database, which are slightly greater than those reported by Niu et al. (2018) (20.1 g kg^{-1} DMI and 6.0%) using an intercontinental dairy cattle database. Ribeiro et al. (2020) and Benaouda et al. (2020) reported 20.9 g kg^{-1} DMI and 6.2%, and 22.9 g kg^{-1} DMI and 7.0% using smaller databases from Brazil and LAC region, respectively. In a concomitant effort

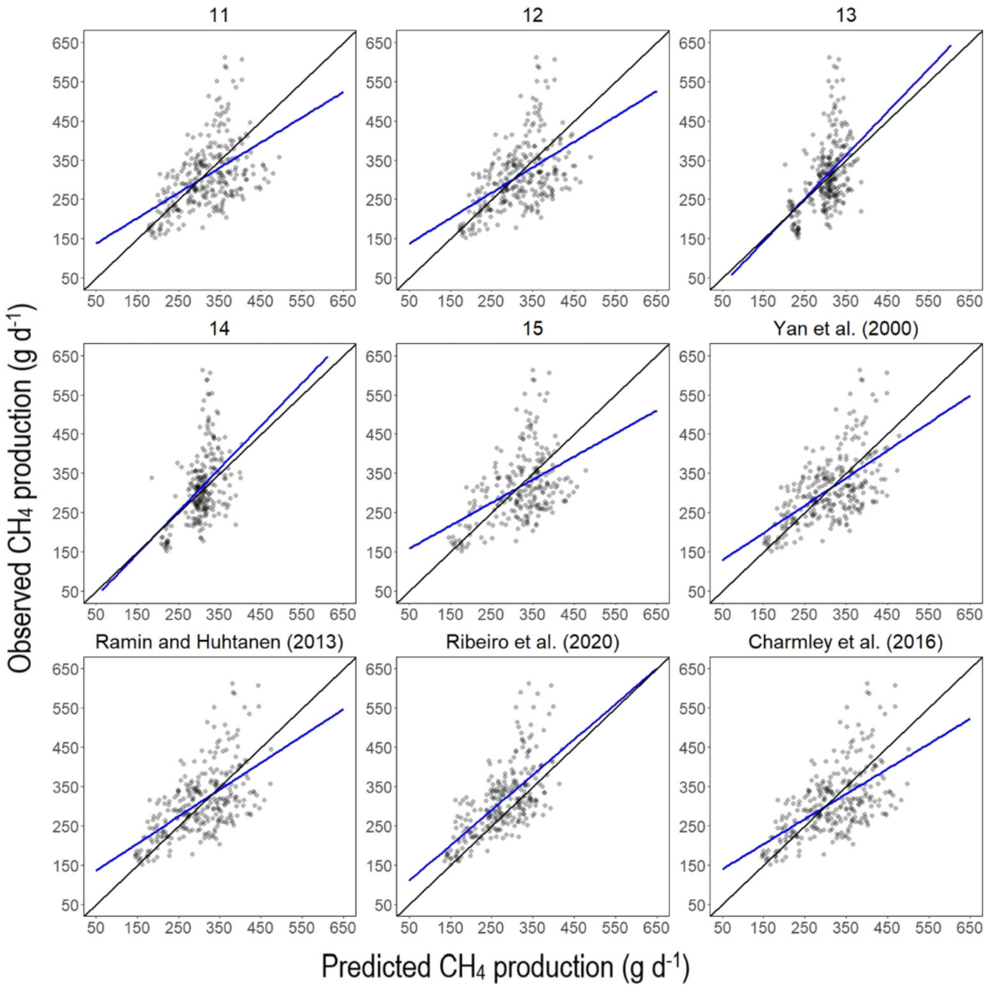


Fig. 3. Observed *vs.* predicted plots for grazing cows methane production (g d^{-1}) prediction equations. Developed models and extant equations are in accordance with Tables 2 and 5, respectively. The blue and black solid lines represent the fitted regression line for the relationship between observed and predicted values, and the identity line (y = x), respectively.

related to the current study, Congio et al. (2022a) reported CH_4 yield and Y_m of LAC beef cattle averaging 22.1 g kg^{-1} DMI and 7.0%, respectively. The CH_4 intensity averaged 21.9 g kg^{-1} MY in the current study, which is greater than 13.7, 17.9, and 19.9 g kg^{-1} MY reported by Niu et al. (2018), Ribeiro et al. (2020), and Benaouda et al. (2020), respectively. The average CH_4 production was 309 g d^{-1}, whereas Niu et al. (2018), Benaouda et al. (2020), and Ribeiro et al. (2020) reported 369, 337, and 292 g d^{-1}, respectively.

Table 3
Performance of extant equations to predict enteric CH_4 production (g d^{-1}) using the all-data subset from the Latin America and Caribbean dairy cattle database (ranked by RSR).

Rank	Reference	Equation[a]	n[b]	RSR[b]	RMSPE, %[b]	MB, %[b]	SB, %[b]	CCC[b]
1	Yan et al. (2000)	(3.234 + 0.0547 × GEI) ÷ 0.05565	610	0.62	19.7	1.33	0.23	0.77
2	Ramin and Huhtanen (2013)	(62 + 25 × DMI) × 0.714	610	0.62	19.9	0.20	0.28	0.77
3	Charmley et al. (2016)	38 + 19.22 × DMI	610	0.64	20.4	3.02	2.25	0.77
4	IPCC (2006)	(0.065 × GEI) ÷ 0.05565	610	0.64	20.5	0.17	8.26	0.79
5	IPCC (1997)	(0.060 × GEI) ÷ 0.05565	610	0.68	21.6	15.33	2.60	0.76
6	Hristov et al. (2013)	2.54 + 19.14 × DMI	610	0.68	21.7	14.70	1.83	0.75
7	Nielsen et al. (2013)	(1.26 × DMI) ÷ 0.05565	610	0.71	22.7	10.84	12.66	0.76
8	Storlein et al. (2014)	(− 1.47 + 1.28 × DMI) ÷ 0.05565	610	0.73	23.3	13.93	13.61	0.76
9	Ellis et al. (2007)	(3.23 + 0.809 × DMI) ÷ 0.05565	610	0.74	23.6	26.69	2.60	0.66
10	Ribeiro et al. (2020)	(0.734 + 0.041 × GEI + 0.009 × BW - 0.04 × EE) ÷ 0.05565	362	0.75	22.4	20.22	0.02	0.66
11	Moraes et al. (2014)	(3.247 + 0.043 × GEI) ÷ 0.05565	610	0.79	25.4	37.19	3.39	0.62
12	Benaouda et al. (2020)	17.0 × DMI + 0.03 × NDF	476	0.82	27.6	44.52	0.00	0.65
13	Patra (2017)	(1.29 + 0.878 × DMI) ÷ 0.05565	610	0.83	26.6	43.79	0.56	0.62
14	Niu et al. (2018)	33.2 + 16.6 × DMI + 2.43 × NDF	574	0.97	31.8	49.19	0.08	0.52

[a] GEI = gross energy intake (MJ d^{-1}); DMI = dry matter intake (kg d^{-1}); BW = body weight (kg); EE = dietary ether extract (% DM); NDF = dietary neutral-detergent fiber (% DM).

[b] *n* = number of observations used for model evaluation; RSR = RMSPE-observations standard deviation ratio; RMSPE = root mean square prediction error (% observed CH_4 production means); MB = mean bias (% MSPE); SB = slope bias (% MSPE); CCC = concordance correlation coefficient.

Table 4
Performance of extant equations to predict enteric CH_4 production (g d^{-1}) using the confined subset from the Latin America and Caribbean dairy cattle database (ranked by RSR).

Rank	Reference	Equation[a]	n[b]	RSR[b]	RMSPE, %[b]	MB, %[b]	SB, %[b]	CCC[b]
1	IPCC (2006)	(0.065 × GEI) ÷ 0.05565	330	0.42	14.7	1.31	0.66	0.91
2	Ramin and Huhtanen (2013)	(62 + 25 × DMI) × 0.714	330	0.45	15.6	0.04	6.43	0.88
3	Yan et al. (2000)	(3.234 + 0.0547 × GEI) ÷ 0.05565	330	0.45	15.7	7.95	6.64	0.88
4	IPCC (1997)	(0.060 × GEI) ÷ 0.05565	330	0.45	15.8	15.03	0.73	0.89
5	Charmley et al. (2016)	38 + 19.22 × DMI	330	0.46	16.0	9.10	1.34	0.88
6	Hristov et al. (2013)	2.54 + 19.14 × DMI	330	0.48	16.8	17.73	1.38	0.87
7	Nielsen et al. (2013)	(1.26 × DMI) ÷ 0.05565	330	0.51	17.9	24.95	3.72	0.87
8	Storlein et al. (2014)	(− 1.47 + 1.28 × DMI) ÷ 0.05565	330	0.53	18.7	29.81	4.62	0.87
9	Ellis et al. (2007)	(3.23 + 0.809 × DMI) ÷ 0.05565	330	0.60	20.9	28.73	19.00	0.77
10	Ribeiro et al. (2020)	(0.734 + 0.041 × GEI + 0.009 × BW - 0.04 × EE) ÷ 0.05565	99	0.60	18.5	10.17	24.81	0.73
11	Moraes et al. (2014)	(3.247 + 0.043 × GEI) ÷ 0.05565	330	0.64	22.4	37.91	19.99	0.74
12	Benaouda et al. (2020)	17.0 × DMI + 0.03 × NDF	256	0.66	23.9	57.27	6.04	0.76
13	Patra (2017)	(1.29 + 0.878 × DMI) ÷ 0.05565	330	0.67	23.5	49.46	9.32	0.74
14	Niu et al. (2018)	33.2 + 16.6 × DMI + 2.43 × NDF	330	0.78	27.3	54.02	8.33	0.65

[a] GEI = gross energy intake (MJ d^{-1}); DMI = dry matter intake (kg d^{-1}); BW = body weight (kg); EE = dietary ether extract (% DM); NDF = dietary neutral-detergent fiber (% DM).
[b] n = number of observations used for model evaluation; RSR = RMSPE-observations standard deviation ratio; RMSPE = root mean square prediction error (% observed CH_4 production means); MB = mean bias (% MSPE); SB = slope bias (% MSPE); CCC = concordance correlation coefficient.

Table 5
Performance of extant equations to predict enteric CH_4 production (g d^{-1}) using the grazing subset from the Latin America and Caribbean dairy cattle database (ranked by RSR).

Rank	Reference	Equation[a]	n[b]	RSR[b]	RMSPE, %[b]	MB, %[b]	SB, %[b]	CCC[b]
1	Yan et al. (2000)	(3.234 + 0.0547 × GEI) ÷ 0.05565	280	0.84	23.6	0.01	9.48	0.59
2	Ramin and Huhtanen (2013)	(62 + 25 × DMI) × 0.714	280	0.85	23.9	0.92	10.18	0.58
3	Ribeiro et al. (2020)	(0.734 + 0.041 × GEI + 0.009 × BW - 0.04 × EE) ÷ 0.05565	263	0.86	24.0	24.49	0.75	0.57
4	Charmley et al. (2016)	38 + 19.22 × DMI	280	0.87	24.6	0.69	14.92	0.59
5	IPCC (2006)	(0.065 × GEI) ÷ 0.05565	280	0.91	25.6	2.19	21.20	0.59
6	Hristov et al. (2013)	2.54 + 19.14 × DMI	280	0.94	26.3	13.84	12.70	0.56
7	Ellis et al. (2007)	(3.23 + 0.809 × DMI) ÷ 0.05565	280	0.94	26.4	25.57	1.39	0.49
8	IPCC (1997)	(0.060 × GEI) ÷ 0.05565	280	0.96	26.8	17.38	12.78	0.55
9	Nielsen et al. (2013)	(1.26 × DMI) ÷ 0.05565	280	0.97	27.2	4.45	26.97	0.58
10	Storlein et al. (2014)	(− 1.47 + 1.28 × DMI) ÷ 0.05565	280	0.99	27.8	6.26	27.78	0.57
11	Moraes et al. (2014)	(3.247 + 0.043 × GEI) ÷ 0.05565	280	1.02	28.5	37.62	0.54	0.44
12	Benaouda et al. (2020)	17.0 × DMI + 0.03 × NDF	220	1.05	31.6	35.60	6.05	0.48
13	Patra (2017)	(1.29 + 0.878 × DMI) ÷ 0.05565	280	1.06	29.8	39.98	2.60	0.45
14	Niu et al. (2018)	33.2 + 16.6 × DMI + 2.43 × NDF	244	1.24	36.9	46.68	4.76	0.33

[a] GEI = gross energy intake (MJ d^{-1}); DMI = dry matter intake (kg d^{-1}); BW = body weight (kg); EE = dietary ether extract (% DM); NDF = dietary neutral-detergent fiber (% DM).
[b] n = number of observations used for model evaluation; RSR = RMSPE-observations standard deviation ratio; RMSPE = root mean square prediction error (% observed CH_4 production means); MB = mean bias (% MSPE); SB = slope bias (% MSPE); CCC = concordance correlation coefficient.

Our study corroborated that feed intake is the key variable predicting CH_4 emission, which agrees with previous reports for dairy (Niu et al., 2018) and beef (van Lingen et al., 2019; Congio et al., 2022a) cattle, goats (Patra and Lalhriatpuii, 2016), and sheep (Patra et al., 2016; Congio et al., 2022b). This relationship clearly indicates that methanogenesis increases when more substrate is available for microbial fermentation in the

Table 6
Enteric CH_4 yield (g kg^{-1} DMI) prediction equations and model performance for the all-data, confined, and grazing subsets of the Latin America and Caribbean dairy cattle database.

Subset	Prediction equation[a]	Model performance[b]					
Equation		n[b]	RSR	RMSPE, %	MB, %	SB, %	CCC
All-data							
(16)	29.5 (1.71) − 0.012 (0.002) × BW − 1.33 (0.296) × FL + 0.102 (0.034) × EPCM	487	0.94	19.8	0.11	2.71	0.15
Confined							
(17)	− 19.8 (14.2) − 0.006 (0.003) × BW + 2.42 (0.780) × GE	247	0.99	13.0	3.04	0.57	0.13
Grazing							
(18)	26.6 (1.55) − 1.25 (0.347) × FL	280	0.98	24.6	1.56	0.23	0.07
(19)	35.8 (2.57) − 0.015 (0.004) × BW − 1.81 (0.359) × FL	280	0.92	23.1	1.24	0.59	0.24

[a] BW = body weight (kg); FL = feeding level (DMI as % BW); MY = milk yield (kg d^{-1}); EPCM = energy and protein-corrected milk (kg d^{-1}); GE = dietary gross energy (MJ kg^{-1} DM).
[b] n = number of observations used to fit equations and for model evaluation; RSR = RMSPE-observations standard deviation ratio; RMSPE = root mean square prediction error (% observed CH_4 production means); MB = mean bias (% MSPE); SB = slope bias (% MSPE); CCC = concordance correlation coefficient.

rumen. Voluntary DMI is a suitable predictor of enteric CH_4 emissions because it is a product of both plant and animal characteristics affecting digestion (Charmley et al., 2016). In the current study, both DMI and GEI were significantly and positively related to CH_4 production with slopes averaging 18.0 g CH_4 kg^{-1} DMI and 1.00 g CH_4 MJ^{-1} GEI, respectively. Still, multiple regression models based only on dietary parameters had the worst predictive performance in all subsets, which aligns with beef (van Lingen et al., 2019; Congio et al., 2022a) and dairy (Niu et al., 2018) cattle previous analyses, and reaffirms the importance of feed intake relative to other predictor variables.

The positive relationship between MY and EPCM with CH_4 emission agrees with Niu et al. (2018), and their eqs. 7 and 8 had prediction ability similar to our MY and EPCM all-data and confined models. It is due to the overall positive relationship between MY and DMI (Niu et al., 2018). Dietary forage content is positively related to CH_4 production (equation not shown), aligning with previous results (Ellis et al., 2007; van Lingen et al., 2019; Benaouda et al., 2020). Increased forage proportion is usually linked with greater NDF concentration in the diet, commonly leading to more acetate and butyrate production, resulting in increased ruminal hydrogen and consequently more CH_4 production (Bannink et al., 2008). The positive relationship between CH_4 production and BW (equation not shown) also agrees with previous research (Moraes et al., 2014; Niu et al., 2018; Benaouda et al., 2020). Rumen volume and BW are proportional and, consequently, heavier animals with higher maintenance energy requirements, tend to ingest more feed and produce more enteric CH_4 (Demment and Van Soest, 1985; Hristov et al., 2013).

Overall, the predictive ability of CH_4 production models increased with model complexity, which aligns with previous studies (Niu et al., 2018; van Lingen et al., 2019). Nevertheless, adding dietary parameters to either DMI or GEI did not increase predictive ability compared with single-based models in all subsets. There was expected at least one dietary parameter being selected with DMI, resulting in a more accurate model similarly reported by previous meta-analyses (Niu et al., 2018; van Lingen et al., 2019). Ribeiro et al. (2020) also found lack of equations including dietary parameters among the best-developed using a lactating dairy cattle subset from Brazil, which overlapped 41% of the current database. This is likely associated with a low variation in diets in both databases, composed of a significant number of feed efficiency trials with a narrow range of dietary nutrient concentrations (Ribeiro et al., 2020). Thus, exploring more contrasting diets to develop more accurate models for the LAC region is recommended in the near future (Congio et al., 2022a). On the other hand, multiple regressions which allowed a selection of all potential predictors had the smallest RMSPE in the all-data and confined subsets. More complex models may have greater applicability in medium- to high-technology dairy systems, where data collection and resources for analysis are available. For low-technology or livelihood farming systems, however, which are typical in the LAC region, simpler models can be more practicable. Under LAC conditions, even DMI may be a restricted variable; thus, models including easily available on-farm covariates (*e.g.*, MY and EPCM) will be more useful (Congio et al., 2022a).

The best CH_4 production developed models had similar predictive performance to the highest-ranked extant equations in all subsets. Overall, simple equations based either on DMI or GEI proposed by Yan et al. (2000), Ramin and Huhtanen (2013), and Charmley et al. (2016) were among the best extant equations in all subsets. The developed model 2, including GEI, had similar predictive performance compared to the first- (Yan et al., 2000) and second- (Ramin and Huhtanen, 2013) ranked extant equations and can be an option for predicting CH_4 production considering all-data. The IPCC (2006) Tier 2 was the best extant equation predicting enteric CH_4 for confined dairy cows. This is probably because the Y_m in IPCC (2006) Tier 2 (6.5%) was close to the average Y_m from confined cows (6.4%) in the current database. Model 7 developed on confined subset was of comparable performance than IPCC (2006) Tier 2 equation and also can be used in this condition. For dairy cows under grazing, model 13, including only MY, had similar predictive performance to those proposed by Yan et al. (2000) and Ramin and Huhtanen (2013), which were first- and second-ranked among grazing extant equations. However, those extant equations over-predicted CH_4 at the high end and under-predicted it at the low end of production. Still, considering that DMI and GEI are not always available in commercial dairy farms, models based on MY or EPCM, which are generally available parameters, can also be used to predict CH_4 production for confined cows.

The predictive ability of CH_4 yield models increased with model complexity, which also agrees with previous research (Niu et al., 2018; van Lingen et al., 2019). Previously mentioned meta-analyses regarding modeling CH_4 emissions focused on CH_4 production while little attention was given to CH_4 yield. Few studies developed CH_4 yield prediction models recently (Niu et al., 2018; van Lingen et al., 2019; Congio et al., 2022a, 2022b). Model evaluations across various complexity levels showed that CH_4 yield of dairy cows under LAC conditions could be predicted reasonably. The developed CH_4 yield models were associated with larger RSR than CH_4 production models, which agrees with Niu et al. (2018) and van Lingen et al. (2019). Body weight, FL, EPCM, and dietary GE were selected for predicting CH_4 yield from dairy cows using the LAC database. The four developed models had RSR < 1.00 and outperformed the equation proposed by Niu et al. (2018) in all conditions.

Research involving enteric CH_4 emissions is relatively recent in the LAC region, and additional research would considerably improve the predictive ability of the present models. Future studies should present a more complete nutrient characterization of the diets, avoiding the need to use literature table values to complete missing parameters in databases (Congio et al., 2022a). Still, a standardization of laboratory procedures (*e.g.*, network trials) by country or region might also be considered. Understanding that most dairy operations in the LAC region are pasture-based, the SF_6 tracer technique is the main CH_4 measurement technique used by LAC researchers and method standardization is highly recommended (Hristov et al., 2018; Jonker et al., 2020; Della Rosa et al., 2021). A standardization of DMI estimation using markers is equally advised (De Souza et al., 2015; Hellwing et al., 2015). Finally, recent research has reported that including digestibility parameters as predictor covariates can increase the overall predictive ability of CH_4 production equations (Benaouda et al., 2020; Ribeiro et al., 2020). However, these models may have limited use in supporting LAC national inventories due to limited availability of those parameters at the farm level.

5. Conclusions

The present analysis is the most comprehensive effort to date to develop enteric CH_4 prediction models for dairy cattle in the LAC region. Feed intake was the primary predictor of CH_4 production, whereas BW and FL were most important in predicting CH_4 yield. Our best-developed CH_4 production models were more accurate than IPCC Tier 2 equations in the all-data and grazing subsets, whereas they had a similar performance for confined dairy systems. Simple regression models containing either MY or EPCM were also accurate in predicting CH_4 production and can be a practical alternative when DMI data are missing. The best-developed CH_4 yield models had a satisfactory accuracy and outperformed extant equations in all subsets. The developed models can be used by police-makers supporting improvements of GHG inventories from LAC countries, which are still based on IPCC equations.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.153982.

Data availability

Individual animal data used in the analysis can be requested by contacting individual contributors which are co-authors on the manuscript.

CRediT authorship contribution statement

Guilhermo F.S. Congio: Investigation, Methodology, Data curation, Formal analysis, Writing – original draft, Visualization, Writing – review

& editing. **André Bannink:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Funding acquisition, Project administration. **Olga L. Mayorga:** Supervision, Project administration. **João P.P. Rodrigues:** Methodology, Formal analysis, Software, Visualization. **Adeline Bougouin:** Methodology, Formal analysis, Software, Visualization. **Ermias Kebreab:** Methodology, Formal analysis, Software, Visualization. **Ricardo R. Silva:** Investigation, Data curation, Writing – review & editing. **Rogério M. Maurício:** Investigation, Data curation, Writing – review & editing. **Sila C. da Silva:** Investigation, Data curation, Writing – review & editing. **Patrícia P.A. Oliveira:** Investigation, Data curation, Writing – review & editing. **Camila Muñoz:** Investigation, Data curation, Writing – review & editing. **Luiz G.R. Pereira:** Investigation, Data curation, Writing – review & editing. **Carlos Gómez:** Investigation, Data curation, Writing – review & editing. **Claudia Ariza-Nieto:** Investigation, Data curation, Writing – review & editing. **Henrique M.N. Ribeiro-Filho:** Investigation, Data curation, Writing – review & editing. **Octavio A. Castelán-Ortega:** Investigation, Data curation, Writing – review & editing. **Jaime R. Rosero-Noguera:** Investigation, Data curation, Writing – review & editing. **Maria Paz:** Investigation, Data curation, Writing – review & editing. **Paulo H.M. Rodrigues:** Investigation, Data curation, Writing – review & editing. **Marcos I. Marcondes:** Investigation, Data curation, Writing – review & editing. **Laura Astigarraga:** Investigation, Data curation, Writing – review & editing. **Sergio Abarca:** Investigation, Data curation, Writing – review & editing. **Alexander N. Hristov:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The LAC methane project was supported by AgResearch (S7-SOW21-Feed/Methane) which was funded by the New Zealand Government to support the objectives of the Livestock Research Group of the Global Research Alliance on Agricultural Greenhouse Gases. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The funding sources that allowed the collaborators to carry out their projects are listed in the Supplementary Material.

References

Arndt, C., Hristov, A.N., Price, W.J., McClelland, S.C., Pelaez, A.M., Cueva, S.F., Oh, J., Bannink, A., Bayat, A.R., Crompton, L.A., Dijkstra, J., Eugène, M.A., Kebreab, E., Kreuzer, M., McGee, M., Martin, C., Newbold, C.J., Reynolds, C.K., Schwarm, A., Shingfield, K.J., Veneman, J.B., Yáñez-Ruiz, D.R., Yu, Z., 2021. Strategies to Mitigate Enteric Methane Emissions by Ruminants – A Way to Approach the 2.0°C Target. CABI preprint 20210085288, Wallingford, UK https://doi.org/10.31220/agriRxiv.2021.00040.

Bannink, A., France, J., Lopez, S., Gerrits, W.J.J., Kebreab, E., Tamminga, S., Dijkstra, J., 2008. Modelling the implications of feeding strategy on rumen fermentation and functioning of the rumen wall. Anim. Feed Sci. Technol. 143, 3–26. https://doi.org/10.1016/j.anifeedsci.2007.05.002.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–51. https://doi.org/10.18637/jss.v067.i01.

Benaouda, M., Martin, C., Li, X., Kebreab, E., Hristov, A.N., Yu, Z., Yáñez-Ruiz, D.R., Reynolds, C.K., Crompton, L.A., Dijkstra, J., Bannink, A., Schwarm, A., Kreuzer, M., McGee, M., Lund, P., Hellwing, A.L.F., Weisbjerg, M.R., Moate, P.J., Bayat, A.R., Shingfield, K.J., Peiren, N., Eugène, M., 2019. Evaluation of the performance of existing mathematical models predicting enteric methane emissions from ruminants: animal categories and dietary mitigation strategies. Anim. Feed Sci. Technol. 114207. https://doi.org/10.1016/J.ANIFEEDSCI.2019.114207.

Benaouda, M., González-Ronquillo, M., Appuhamy, J., Kebreab, E., Molina, L.T., Herrera-Camacho, J., Ku-Vera, J.C., Ángeles-Hernández, J.C., Castelán-Ortega, O., 2020. Development of mathematical models to predict enteric methane emission by cattle in Latin America. Livest. Sci. 241, 104177. https://doi.org/10.1016/j.livsci.2020.104177.

Bibby, J., Toutenburg, T., 1977. Prediction and Improved Estimation in Linear Models. John 785 Wiley Sons, Chichester.

Charmley, E., Williams, S.R.O., Moate, P.J., Hegarty, R.S., Herd, R.M., Oddy, V.H., Reyenga, P., Staunton, K.M., Anderson, A., Hannah, M.C., 2016. A universal equation to predict methane production of forage-fed cattle in Australia. Anim. Prod. Sci. 56, 169–180. https://doi.org/10.1071/AN1536.

Congio, G.F.S., Bannink, A., Mayorga-Mogollón, O.L., Latin America Methane Project Collaborators, Hristov, A.N., 2021. Enteric methane mitigation strategies for ruminant livestock systems in the Latin America and Caribbean region: a meta-analysis. J. Clean. Prod. 312, 127693. https://doi.org/10.1016/j.jclepro.2021.127693.

Congio, G.F.S., Bannink, A., Mayorga, O.L., Rodrigues, J.P.P., Bougouin, A., Kebreab, E., Carvalho, P.C.F., Berchielli, T.T., Mercadante, M.E.Z., Valadares-Filho, S.C., Borges, A.L.C.C., Berndt, A., Rodrigues, P.H.M., Ku-Vera, J.C., Molina-Botero, I.C., Arango, J., Reis, R.A., Posada-Ochoa, S.L., Tomich, T.R., Castelán-Ortega, O.A., Marcondes, M.I., Gómez, C., Ribeiro-Filho, H.M.N., Gere, J.I., Ariza-Nieto, C., Giraldo, L.A., Gonda, H., Cerón-Cucchi, M.E., Hernández, O., Ricci, P., Hristov, A.N., 2022a. Prediction of enteric methane production and yield in beef cattle using a Latin America and Caribbean database (Under Review).

Congio, G.F.S., Bannink, A., Mayorga, O.L., Rodrigues, J.P.P., Bougouin, A., Kebreab, E., Carvalho, P.C.F., Abdalla, A.L., Monteiro, A.L.G., Ku-Vera, J.C., Gere, J.I., Gómez, C., Hristov, A.N., 2022b. Prediction of enteric methane production and yield in sheep cattle using a Latin America and Caribbean database (Under Review).

De Souza, J., Batistel, F., Welter, K.C., Silva, M.M.V., Costa, D.F.A., Santos, F.A.P., 2015. Evaluation of external markers to estimate fecal excretion, intake and digestibility in dairy cows. Trop. Anim. Health Prod. 47, 265–268. https://doi.org/10.1007/s11250-014-0674-6.

Della Rosa, M.M., Jonker, A., Waghorn, G.C., 2021. A review of technical variations and protocols used to measure methane emissions from ruminants using respiration chambers, SF6 tracer technique and GreenFeed, to facilitate global integration of published data. Anim. Feed Sci. Technol. 279, 115018. https://doi.org/10.1016/j.anifeedsci.2021.115018.

Demment, M.W., Van Soest, P.J., 1985. A nutritional explanation for body-size patterns of ruminant and non-ruminant herbivores. Am. Nat. 125, 641–672. https://doi.org/10.1086/284369.

Ellis, J.L., Kebreab, E., Odongo, N.E., McBride, B.W., Okine, E.K., France, J., 2007. Prediction of methane production from dairy and beef cattle. J. Dairy Sci. 90, 3456–3467. https://doi.org/10.3168/jds.2006-675.

FAOSTAT, 2020. FAO statistical database. Food and Agricultural Organization of the United Nations, Rome (accessed on 23 April 2021) http://www.fao.org/faostat/en/#data.

Hammond, K.J., Crompton, L.A., Bannink, A., Dijkstra, J., Yáñez-Ruiz, D.R., O'Kiely, P., Kebreab, E., Eugène, M.A., Yu, Z., Shingfield, K.J., Schwarm, A., Hristov, A.N., Reynolds, C.K., 2016. Review of current in vivo measurement techniques for quantifying enteric methane emission from ruminants. Anim. Feed Sci. Technol. 219, 13–30. https://doi.org/10.1016/j.anifeedsci.2016.05.018.

Hellwing, A.L.F., Lund, P., Weisbjerg, M.R., Oudshoorn, F.W., Munksgaard, L., Kristensen, T., 2015. Comparison of methods for estimating herbage intake in grazing dairy cows. Livest. Sci. 176, 61–74. https://doi.org/10.1016/j.livsci.2015.01.013.

Hristov, A.N., Oh, J., Firkins, J., Dijkstra, J., Kebreab, E., Waghorn, G., Makker, M.P.S., Adesogan, A.T., Yang, W., Lee, C., Gerber, P.J., Henderson, B., Tricarico, J.M., 2013. SPECIAL TOPICS - mitigation of methane and nitrous oxide emissions from animal operations: I. A review of enteric methane mitigation options. J. Anim. Sci. 91, 5045–5069. https://doi.org/10.2527/jas.2013-6583.

Hristov, A.N., Kebreab, E., Niu, M., Oh, J., Bannink, A., Bayat, A.R., Boland, T., Brito, A.F., Casper, D., Crompton, L.A., Dijkstra, J., Eugène, M.A., Garnsworthy, Ph.C., Haque, M.N., Hellwing, A.L.F., Huhtanen, P., Kreuzer, M., Kuhla, B., Lund, P., Madsen, J., Martin, C., Moate, P.J., Muetzel, S., Muñoz, C., Peiren, N., Powell, J.M., Reynolds, C.K., Schwarm, A., Shingfield, K.J., Storlien, T.M., Weisbjerg, M.R., Yáñez-Ruiz, D.R., Yu, Z., 2018. Symposium review: uncertainties in enteric methane inventories, measurement techniques, and prediction models. J. Dairy Sci. 101, 6655–6674. https://doi.org/10.3168/jds.2017-13536.

IPCC, 1997. Revised 1996 IPCC guidelines for national greenhouse gas inventories. Intergovernmental Panel on Climate Change, IPCC/OECD/IEA, Bracknell, UK (accessed on 26 April 2021) https://www.ipcc-nggip.iges.or.jp/public/gl/invs1.html.

IPCC, 2006. 2006 IPCC guidelines for national greenhouse gas inventories. Intergovernmental Panel on Climate Change, IGES, Kanagawa, Japan (accessed on 26 April 2021) https://www.ipcc-nggip.iges.or.jp/public/2006gl/.

IPCC, 2018. In: Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J.B.R., Chen, Y., Zhou, X., Gomis, M.I., Lonnoy, E., Maycock, T., Tignor, M., Waterfield, T. (Eds.), An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty . https://www.ipcc.ch/sr15/. (Accessed 20 October 2021).

James, G., Witten, D., Hastie, T., Tibshirani, R., 2014. An Introduction to Statistical Learning: With Applications in R. Springer, New York, US.

Jonker, A., Waghorn, G., Berndt, A., Boland, T., Deighton, M.H., Gere, J.I., Grainger, C., Hegarty, R.S., Iwaasa, A.D., Koolaard, J.P., Lassey, K.R., Luo, D., Martin, C., Moate, P.J., Molano, G., Pinares-Patiño, C.S., Ribaux, B.E., Rochette, Y., Silvestre, T.M., Swainson, N.M., van Wyngaard, J.D.V., Williams, R.S.O., 2020. Guidelines for use of sulphur hexafloride (SF6) tracer technique to measure enteric methane emissions from ruminants. New Zealand Agricultural Greenhouse Gas Research Centre, New Zealand (accessed on 24 May 2021) https://globalresearchalliance.org/publication-library/?doctype=223&research-group=212.

Lin, L.I., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268. https://doi.org/10.2307/2532051.

Moraes, L.E., Strathe, A.B., Fadel, J.G., Casper, D.P., Kebreab, E., 2014. Prediction of enteric methane emissions from cattle. Glob. Change Biol. 20, 2140–2148. https://doi.org/10.1111/gcb.12471.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153.

Nielsen, N.I., Volden, H., Akerlind, M., Brask, M., Hellwing, A.L.F., Storlein, T., Bertilsson, J., 2013. A prediction equation for enteric methane emission from dairy cows for use in NorFor. Acta Agric. Scand. Sect. A-Anim. Sci. 63, 126–130. https://doi.org/10.1080/09064702.2013.851275.

Niu, M., Kebreab, E., Hristov, A.N., Oh, J., Arndt, C., Bannink, A., Bayat, A.R., Brito, A.F., Boland, T., Casper, D., Crompton, L.A., Dijkstra, J., Eugène, M.A., Garnsworthy, P.C., Haque, M.N., Hellwing, A.L.F., Huhtanen, P., Kreuzer, M., Kuhla, B., Lund, P., Madsen, J., Martin, C., McClelland, S.C., McGee, M., Moate, P.J., Muetzel, S., Muñoz, C., O'Kiely, P., Peiren, N., Reynolds, C.K., Schwarm, A., Shingfield, K.J., Storlien, T.M., Weisbjerg, M.R., Yáñez-Ruiz, D.R., Yu, Z., 2018. Prediction of enteric methane production, yield, and intensity in dairy cattle using an intercontinental database. Glob. Change Biol. 24, 3368–3389. https://doi.org/10.1111/gcb.14094.

NRC, 2001. Nutrient Requirements of Dairy Cattle. 7th revised edition. National Research Council, National Academy Press, Washington, DC.

Patra, A.K., 2017. Prediction of enteric methane emission from cattle using linear and non-linear statistical models in tropical production systems. Mitig. Adapt. Strateg. Glob. Change 22, 629–650. https://doi.org/10.1007/s11027-015-9691-7.

Patra, A.K., Lalhriatpuii, M., 2016. Development of statistical models for prediction of enteric methane emission from goats using nutrient composition and intake variables. Agric. Ecosyst. Environ. 215, 89–99. https://doi.org/10.1016/j.agee.2015.09.018.

Patra, A.K., Lalhriatpuii, M., Debnath, B.C., 2016. Predicting enteric methane emission in sheep using linear and non-linear statistical models from dietary variables. Anim. Prod. Sci. 56, 574–584. https://doi.org/10.1071/AN15505.

R Core Team, 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (accessed on 20 Sep 2021) https://www.R-project.org/.

Ramin, M., Huhtanen, P., 2013. Development of equations for predicting methane emissions from ruminants. J. Dairy Sci. 96, 2476–2493. https://doi.org/10.3168/jds.2012-6095.

Ribeiro, R.S., Rodrigues, J.P.P., Mauricio, R.M., Borges, A.L.C.C., Berchielli, T.T., Valadares-Filho, S.C., Machado, F.S., Campos, M.M., Tomich, T.R., Pereira, L.G.R., 2020. Predicting enteric methane production from cattle in the tropics. Animal 14, s438–s452. https://doi.org/10.1017/S1751731120001743.

Storlein, T.M., Volden, H., Almoy, T., Beauchemin, K.A., Mcallister, T.A., Harstad, O.M., 2014. Prediction of enteric methane production from dairy cows. Acta Agric. Scand. Sect. A-Anim. Sci. 64, 98–109. https://doi.org/10.1080/09064702.2014.959553.

UN General Assembly, 2015. Transforming our world: the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1. https://www.refworld.org/docid/57b6e3e44.html (accessed on 26 April 2021).

van Lingen, H.J., Niu, M., Kebreab, E., Valadares-Filho, S.C., Rooke, J.A., Duthie, C.-A., Schwarm, A., Kreuzer, M., Hynd, P.I., Caetano, M., Eugène, M., Martin, C., McGee, M., O'Kiely, P., Hünerberg, M., McAllister, T.A., Berchielli, T.T., Messana, J.D., Peiren, N., Chaves, A.V., Charmley, E., Cole, N.A., Hales, K.E., Lee, S.-S., Berndt, A., Reynolds, C.K., Crompton, L.A., Bayat, A.-R., Yáñez-Ruiz, D.R., Yu, Z., Bannink, A., Dijkstra, J., Casper, D.P., Hristov, A.N., 2019. Prediction of enteric methane production, yield and intensity of beef cattle using an intercontinental database. Agric. Ecosyst. Environ. 283, 106575. https://doi.org/10.1016/j.agee.2019.106575.

Weiss, W.P., Tebbe, A.W., 2019. Estimating digestible energy values of feeds and diets and integrating those values into net energy systems. TAS 3 (3), 953–961. https://doi.org/10.1093/tas/txy119.

Yan, T., Agnew, R.E., Gordon, F.J., Porter, M.G., 2000. Prediction of methane energy output in dairy and beef cattle offered grass silage-based diets. Livest. Prod. Sci. 64, 253–263. https://doi.org/10.1016/S0301-6226(99)00145-1.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x.

Zwillinger, D., Kokoska, S., 2000. CRC Standard Probability and Statistics Tables and Formulae. CRC Press, Boca Raton, US.